PCT

(51) International Patent Classification 5 : C07K 15/06, 3/02, 3/20, 13/00, A61K 35/12 A61K 37/12, A01N 63/02, A23J 1/10	**A1**	**(11) International Publication Number:** **WO 91/02744** **(43) International Publication Date:** 7 March 1991 (07.03.91)

(21) International Application Number: PCT/US90/04745

(22) International Filing Date: 21 August 1990 (21.08.90)

(30) Priority data:
397,779   21 August 1989 (21.08.89)   US

(60) Parent Application or Grant
(63) Related by Continuation
US   397,779 (CIP)
Filed on   21 August 1989 (21.08.89)

(71) Applicant *(for all designated States except US)*: CELTRIX LABORATORIES, INC. [US/US]; 2500 Faber Place, Palo Alto, CA 94303 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)* : BENTZ, Hanne [DE/US]; 36125 Toulouse Street, Newark, CA 94560 (US). NATHAN, Ranga [CA/US]; 6104 Robertson Avenue, Newark, CA 94560 (US). ROSEN, David, M. [US/US]; 3655 Shadyhollow Court, San Jose, CA 95148 (US). DASCH, James, R. [US/US]; 3181 Morris Drive, Palo Alto, CA 94303 (US). SEYEDIN, Saeid, M. [IR/US]; 20761 Lowena Court, Saratoga, CA 95070 (US). PIEZ, Karl, A. [US/US]; 1120 Bay Laurel Drive, Menlo Park, CA 94025 (US). OGAWA, Yasushi [US/US]; 310 Farallon Avenue, Pacifica, CA 94044 (US). ANDREWS, William, H. [US/US]; 780 Fathom Drive, San Mateo, CA 94404 (US). STEPHENSON, Frank, H. [US/US]; 710

Moraga, Apartment C, San Francisco, CA 94122 (US). HEDGPETH, Joel [US/US]; 200 Cresta Vista, San Francisco, CA 94127 (US).

(74) Agents: CIOTTI, Thomas, E. et al.; Irell & Manella, 545 Middlefield Road, Suite 200, Menlo Park, CA 94025 (US).

(81) Designated States: AT (European patent), AU, BE (European patent), CA, CH (European patent), DE (European patent)*, DK (European patent), ES (European patent), FR (European patent), GB (European patent), IT (European patent), JP, LU (European patent), NL (European patent), SE (European patent), US.

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: BONE-SPECIFIC PROTEIN

(H_2N)-Ala-Lys-Tyr-Asn-Lys-Ile-Lys-Ser-Arg-Gly-
Ile-Lys-Ala-Asn-R^1-Phe-Lys-Lys-Leu-R^2-
Asn-Leu-R^3-Phe-Leu-Tyr-Leu-Asp-His-Asn-
Ala-Leu-Glu-Ser-Val-Pro-Leu-Asn-Leu-Pro-
Glu-Ser-Leu-Arg-Val-Ile-His-Leu-Gln-Phe-
Asn-Asn-Ile-R^4-Ser-Ile-Thr-Asp-Asp-Thr-   (I)
Phe-Cys-Lys-Ala-Asn-Asp-Thr-Ser-Tyr-Ile-
Arg-Asp-Arg-Ile-Glu-Glu-Ile-Arg-Leu-Glu-
Gly-Asn-Pro-R^5-R^6-Leu-Gly-Lys-His-Pro-
Asn-Ser-Phe-Ile-Cys-Leu-Lys-Arg-Leu-Pro-
Ile-Gly-Ser-Tyr-R^7-(COOH),

DEMINERALIZED BONE POWDER (0.5 N HCL)

↓

4 M GUANIDINE HYDROCHLORIDE EXTRACT

↓

SEPHACRYLS — 200 (20 — 35KD MW)

↓

CM — CELLULOSE (10 — 150 mM NaCL)

↓

CONCANAVALIN — A (0.5 M METHYL $\propto$ — D — MANNOPYRANOSIDE)

↓

HEPARIN — SEPHAROSE (0.1 — 0.5 M NaCL)

↓

C 18 — REVERSED PHASE — HPLC (42 — 45% ACETONOTRILE)

(57) Abstract

Proteins, having sequence (I) where R^1 is Ala ou Thr, R^2 is Asn or His, R^3 is Thr or Ser, R^4 is Ala or Thr, R^5 is Ile or Val, R^6 is Val or Ile and R^7 is Phe or Ile, is disclosed. These proteins are present in osteoblasts and osteocytes and may be used as markers to follow bone and/or cartilage metabolism.

* See back of page

DESIGNATIONS OF "DE"

Until further notice, any designation of "DE" in any international application whose international filing date is prior to October 3, 1990, shall have effect in the territory of the Federal Republic of Germany with the exception of the territory of the former German Democratic Republic.

-1-

5   BONE-SPECIFIC PROTEIN

                              Description

      Technical Field
10   The present invention relates to protein
      chemistry. More particularly, it relates a protein that
      occurs specifically in bone, antibodies to such proteins,
      and immunoassays for detecting such proteins.

15   Background Art
              Others have fractionated bone in an attempt to
      identify proteins which can stimulate the formation of new
      bone when placed in contact with living systems. (Urist,
      M. R., Clin Orthop (1968) 56:37; Science (1965) 150:893;
20   Reddi, A. H., et al., Proc Natl Acad Sci (USA) (1972)
      69:1601.) A "bone morphogenic protein" (BMP) was
      extracted from demineralized bone using urea or guanidine
      hydrochloride and reprecipitated according to the
      disclosures in U.S. Patents Nos. 4,294,753 and 4,455,256
25   to Urist. Urist subsequently reported (Urist, M. R., Clin
      Orthop Rel Res (1982) 162:219) that ion exchange
      purification of this crude protein mixture yielded an
      activity which was unadsorbed to carboxymethyl cellulose
      resin (CMC) at pH 4.8. Urist's reports in Science (1983)
30   220:680-685, Proc Natl Acad Science (USA) (1984) 81:371-
      375, and U.S. Pat. No. 4,789,732 describe BMPs having
      molecular weights of 17,500 and 18,500 daltons. Urist's
      patent publication, EPA Publication No. 0212474, describes
      BMP fragments of 4,000 to 7,000 daltons obtained by
35   limited proteolysis of BMP.

-2-

U.S. Patent No. 4,608,199 describes a bone-
derived protein of 30,000-32,000 daltons. The protein is
described as being water soluble and having no affinity
for concanavalin A (ConA).

5   WO 88/00205 reports four proteins, designated
BMP-1, BMP-2 Class I, BMP-2 Class II and BMP-3, that are
alleged to have osteogenic activity by themselves or in
combination with other factors. Sequences are provided
for each of these proteins which show no homology to the
10   sequence (see below) of the protein of the present
invention.
           Commonly owned U.S. 4,434,094 reported the
partial purification of a bone generation-stimulating,
bone-derived protein by extraction with chaotropic agents,
15   fractionation on anion and cation exchange columns, and
recovery of the activity from a fraction adsorbed to CMC
at pH 4.8. This new protein fraction was termed
"osteogenic factor" (OF) and was characterized as having a
molecular weight below about 30,000 daltons.
20   Commonly owned U.S. Patent No. 4,774,332
describes two proteins that were purified to homogeneity
using a purification procedure that is similar in part to
that disclosed in U.S. 4,434,094. Those two proteins
eluted from CMC at about a 150-200 mM NaCl gradient.
25   These two proteins were originally called cartilage-
inducing factor (CIF) A and CIF B. CIF A was subsequently
found to be identical to a previously identified protein
now called transforming growth factor beta1 (TGF-beta1).
CIF B has been found to be a novel form of TGF-beta and is
30   now known as TGF-beta2. These proteins and homologous
proteins exhibiting similar activity are collectively
referred to as TGF-beta.
           Commonly owned U.S. Patent No. 4,627,982
concerns a partially purified bone-inducing factor present
35   in the CMC-bound fraction of U.S. 4,434,094 that elutes in
the portion of the NaCl gradient below that in which the

-3-

major portions of TGF-beta1 and TGF-beta2 elute (i.e.,
below about 150 mM NaCl). The present invention relates
to the identification of an ingredient of that fraction.

5 Disclosure of the Invention

One aspect of the invention is a substantially
pure polypeptide that is found in bone and has the
following sequence:
10

(H_2N)-Ala-Lys-Tyr-Asn-Lys-Ile-Lys-Ser-Arg-Gly-
Ile-Lys-Ala-Asn-R^1-Phe-Lys-Lys-Leu-R^2-
Asn-Leu-R^3-Phe-Leu-Tyr-Leu-Asp-His-Asn-
Ala-Leu-Glu-Ser-Val-Pro-Leu-Asn-Leu-Pro-
15   Glu-Ser-Leu-Arg-Val-Ile-His-Leu-Gln-Phe-
Asn-Asn-Ile-R^4-Ser-Ile-Thr-Asp-Asp-Thr-
Phe-Cys-Lys-Ala-Asn-Asp-Thr-Ser-Tyr-Ile-
Arg-Asp-Arg-Ile-Glu-Glu-Ile-Arg-Leu-Glu-
Gly-Asn-Pro-R^5-R^6-Leu-Gly-Lys-His-Pro-
20   Asn-Ser-Phe-Ile-Cys-Leu-Lys-Arg-Leu-Pro-
Ile-Gly-Ser-Tyr-R^7-(COOH),

where R^1 is Ala or Thr, R^2 is Asn or His, R^3 is Thr or
Ser, R^4 is Ala or Thr, R^5 is Ile or Val, R^6 is Val or Ile
25 and R^7 is Phe or Ile, and substantially pure polypeptides
that are and substantially homologous thereto.

Deglycosylated analogs of the above-described
polypeptides are another aspect of the invention.

Further aspects of the invention are recombinant
30 materials (i.e., recombinant DNA, recombinant vectors, and
recombinant cells or microorganisms) and processes for
producing the polypeptides of the invention, antibodies
specific to the polypeptides, and immunoassays for the
polypeptides.

35

-4-

Brief Description of the Drawings .

        In the drawings:

        Figure 1 is a flow chart of the process that was
used to isolate the bovine species of the bone-specific
5 protein of the invention from demineralized bovine bone.

        Figure 2 is a graph of the optical densities
(absorbances at 280 nm) of the gel filtration fractions of
the gel filtration fractions of the example (¶C).

        Figure 3 is a graph of the optical densities
10 (absorbances at 280 nm) of eluate fractions from the
preparative ion exchange chromatography of the example
(¶D).

        Figure 4 is a graph of the optical densities
(absorbances at 280 nm) of eluate fractions from the
15 cross-linked ConA chromatography step of the example (¶E);

        Figure 5 is a graph of the optical densities
(absorbances at 280 nm) of eluate fractions from the
heparin-sepharose chromatography step of the example (¶F);

        Figure 6 is a graph of the optical densities
20 (absorbances at 230 nm) of the gradient fractions from the
C18-RP-HPLC chromatography step of the example (¶G);

        Figure 7 is a table showing results of amino
acid sequencing of the bovine isolate of the invention and
locations of the sequenced fragments in the overall
25 sequence.

        Figure 8 is a photograph of an autoradiograph of
SDS-PAGE analyses of the purified bovine protein that are
described in the example (¶H) (lanes A and C show
glycosylated protein; lanes B and D show enzymatically
30 deglycosylated protein).

        Figure 9 is a schematic diagram illustrating the
structure of the bovine gene that encodes the mature
bovine species of the protein of the invention. Various
restriction sites are indicated.

35

Figure 10 is a restriction map of the region of
the human gene that encodes the human species of the
protein.

Figure 11, in parts A-E, shows the DNA sequence
5   and deduced amino acid sequence of the mature human spe-
cies of the protein of the invention in comparison with
the sequences of the bovine gene and protein. Part A
shows the preliminary unconfirmed sequence of a putative
"amino terminal" precursor exon of the human gene for the
10 prepropolypeptide. The extent of this exon has not been
determined. Bases of uncertain identity in the DNA
sequence are represented by numbers. The signal sequence
indicated in the figure is putative and was identified by
the von Heijne algorithm. Part B shows the sequence of
15 another precursor exon, designated -1, that is downstream
of the amino terminal exon and upstream of the exon that
encodes the amino terminal of the mature protein. Parts
C-E show the sequences of exons 1-3 which encode the
mature protein. Asterisks designate points of identity
20 between the bovine and human DNA sequences. Amino acid
numbering is relative to the mature sequence with the
first amino acid thereof designated "1". Consensus 3' and
5' exon splice sites are indicated by YYYYYYYYYYYNYAG and
JAGGTRAGT, respectively.
25   Figure 12 is a schematic diagram of the
mammalian expression vector phOIF18 which contains the
human gene that encodes the human species of the invention
protein.

30 Modes of Carrying Out the Invention

Isolation of Protein from Bone

It is believed that the protein of the present
invention has been highly conserved among mammalian
35 species--i.e., corresponding proteins from different
mammalian species (herein called "species analogs") will

-6-

have substantially homologous amino acid sequences that
vary from the bovine or human proteins described herein,
if at all, in one or more amino acid residue additions,
deletions or substitutions and/or substantially similar
5   glycosylation patterns. The amino acid sequences of
"substantially homologous" proteins will usually be at
least 50% identical, more usually at least 80% identical,
and preferably at least 90% identical to the bovine/human
amino acid sequence described herein. Such proteins may
10   be derived from bone or other tissues of diverse mammalian
origin or synthesized using recombinant DNA procedures.
The term is intended to include muteins or analogs of the
native protein that are altered in manners known in the
art, such as by substitution of cysteines with neutral
15   (uncharged) amino acids to avoid improper disulfide
bonding, by substitution or elimination of residues in the
asparagine-linked glycosylation sites of the proteins to
alter glycosylation patterns, by substitution of
methionines to make the molecules less susceptible to
20   oxidation, by conservative substitution of other residues,
by chemical modification of one or more residues, by
substitution with nonnatural amino acids or by elimination
or alteration of side-chain sugars. The source of protein
prepared by purification from native sources is
25   advantageously porcine or bovine long bone because of its
ready availability.
         The process for isolating the protein from bone
is as follows. The bone is first cleaned using mechanical
or abrasive techniques, fragmented, and further washed
30   with, for example, dilute aqueous acid preferably at low
temperature. The bone is then demineralized by removal of
the calcium phosphates in their various forms, usually by
extraction with stronger acid. These techniques are
understood in the art, and are disclosed, for example, in
35   U.S. 4,434,094. The resulting preparation, a

demineralized bone, is the starting material for the
preparation of the protein from native sources.

The initial extraction is designed to remove the
nonfibrous (e.g., noncollagenous) proteins from the
5   demineralized bone. This can be done with the use of
chaotropic agents such as guanidine hydrochloride (at
least about 4 molar), urea (8 molar) plus salt, or sodium
dodecylsulfate (at least about 1% by volume) or such other
chaotropic agents as are known in the art (Termine et al.,
10 J Biol Chem (1980) 255:9760-0772; and Sajera and Hascall,
J Biol Chem (1969) 244:77-87 and 2384-2396). The extrac-
tion is preferably carried out at reduced temperatures to
reduce the likelihood of digestion or denaturation of the
extracted protein. A protease inhibitor may be added to
15 the extractant, if desired. The pH of the medium depends
upon the extractant selected. The process of extraction
generally takes on the order of about 4 hr to 1 day.

After extraction, the extractant may be removed
by suitable means such as dialysis against water, preceded
20 by concentration by ultrafiltration if desired. Salts can
also be removed by controlled electrophoresis, or by
molecular sieving, or by any other means known in the art.
It is also preferred to maintain a low temperature during
this process so as to minimize denaturation of the
25 proteins. Alternatively, the extractant chaotropic agent
need not be removed, but rather the solution need only be
concentrated, for example, by ultrafiltration.

The extract, dissolved or redissolved in
chaotropic agent, is subjected to gel filtration to obtain
30 fractions of molecular weight in the range of about 20,000
to 36,000 daltons. Gel sizing is done using standard
techniques, preferably on a Sephacryl S-200 column at room
(10°C-25°C) temperature.

The sized fraction is then subjected to ion
35 exchange chromatography using CMC at approximately pH 4.5-
5.2 preferably about 4.8, in the presence of a nonionic

-8-

chaotropic agent such as 6M urea. Other cation exchangers
may be used, including those derived from polyacrylamide
and cross-linked dextran; however cellulosic cation
exchangers are preferred. Of course, as in any ion
5   exchange procedure, the solution must be freed of
competing ions before application to the column. The
protein is adsorbed on the column and is eluted in an
increasing salt concentration gradient in the range of
about 10 mM to about 150 mM. This fraction is designated
10 "CMB-1" for convenience.

       CMB-1 is lyophilized and the dry CMB-1 is dis-
solved in aqueous sodium deoxycholate (DOC), pH 8.0. This
solution is affinity chromatographed on an equilibrated
column of ConA cross-linked to resin. The ConA-bound
15 material is eluted from the resin with aqueous DOC
containing a displacement carbohydrate. This fraction is
designated "CAB-1" for convenience.

       CAB-1 is reequilibrated for heparin-sepharose
chromatography by desalting on a GH-25 column equilibrated
20 on heparin-sepharose buffer, 6M urea, 0.1M NaCl, 50 mM
Tris-HCl pH 7.2. The desalted fraction is loaded onto a
heparin-sepharose column. After washing, bound material
is eluted from the column using the same buffer at a 0.5M
NaCl salt concentration. The resulting eluate is
25 designated "HSB-1" for convenience.

       HSB-1 is diluted and adjusted to pH 2 and loaded
onto a C18-RP-HPLC column. Bound proteins were gradient
eluted from the column using a solvent consisting of 90%
acetonitrile in 0.1% aqueous TFA (Solvent B). The protein
30 of the invention elutes at approximately 47-50% of solvent
B (42-45% acetonitrile) by volume.

       Proteins eluted by the C18 chromatography were
iodinated by the chloramine-T method. Analysis of the
fraction by SDS-PAGE and autoradiography shows a major
35 broad band at 20,000 to 28,000 daltons comprising the

-9-

protein. The "smearing" of the protein is believed to
mainly be the result of heterogeneity in the glycosylation
of the molecule or perhaps variable post-translational
modification or proteolytic degradation. After enzymatic
5 or chemical deglycosylation, SDS-PAGE analysis of the
protein gives a single band of approximately 10,000
daltons. Reduction of the deglycosylated protein with
dithiothreitol does not affect its migration.
          Initial amino acid sequence analysis of the
10 glycosylated protein so isolated from bovine bone yielded
the following internal sequence in the N-terminal portion
of the protein:

          -Lys-Tyr-Asn-Lys-Ile-Lys-Ser-Arg-Gly-Ile-Lys-
15 Ala-Asn-Thr-Phe-Lys-Lys-Leu-His-Asn-Leu-Ser-Phe-X-Tyr-Thr-
Asp-His-Asn-Ala-Leu-Glu-

The initial amino acid (Lys) in the above sequence is
nearest to the N-terminal. Initially, the nature of the
20 signal obtained for the residue designated X did not
permit this residue to be identified. Repeated sequencing
of the entire peptide and sequencing of oligopeptides
generated from endoproteinase Lys-C (an enzyme that
cleaves proteins at Lys residues) and endoproteinase Glu-C
25 (an enzyme that cleaves proteins at Glu residues) digests
have revealed that the above sequence is preceded by an
Ala residue which is the N-terminus, that the residue
designated X is Leu, that the second Thr residue (the 26th
residue in the above sequence) was incorrect and that this
30 residue is actually a Leu residue, and that the isolate
consists of a protein of approximately 106 amino acids.
Figure 7 provides a summary of these sequence analyses.
The symbol "CHO" designates a carbohydrate substituent.
The symbol "COOH" represents a carboxyl group and
35 designates the carboxy terminus. The first column (on the

-10-

left) provides the sequence analysis of the N-terminal
fragment described above. The second, fourth, and sixth
columns give the sequences of three major Lys-C fragments
of the isolate. The third and fifth columns give the
sequences of two Glu-C fragments.

5   Subsequent isolation of the gene for this
protein confirmed the sequence shown in Figure 7 with the
sole exception that the deduced sequence lacked the Asp
residue at the carboxy terminal. Accordingly, the
sequence for the native bovine protein is as follows:

10

> (H₂N)-Ala-Lys-Tyr-Asn-Lys-Ile-Lys-Ser-Arg-Gly-
>   Ile-Lys-Ala-Asn-Thr-Phe-Lys-Lys-Leu-His-
>   Asn-Leu-Ser-Phe-Leu-Tyr-Leu-Asp-His-Asn-
>   Ala-Leu-Glu-Ser-Val-Pro-Leu-Asn-Leu-Pro-
>   Glu-Ser-Leu-Arg-Val-Ile-His-Leu-Gln-Phe-
>   Asn-Asn-Ile-Thr-Ser-Ile-Thr-Asp-Asp-Thr-
>   Phe-Cys-Lys-Ala-Asn-Asp-Thr-Ser-Tyr-Ile-
>   Arg-Asp-Arg-Ile-Glu-Glu-Ile-Arg-Leu-Glu-
>   Gly-Asn-Pro-Val-Ile-Leu-Gly-Lys-His-Pro-
>   Asn-Ser-Phe-Ile-Cys-Leu-Lys-Arg-Leu-Pro-
>   Ile-Gly-Ser-Tyr-Ile-(COOH),

15

20

The sequence of the corresponding human protein
was determined by obtaining the human gene using DNA

25 probes based on the bovine DNA sequence, sequencing the
human gene and deducing the amino acid sequence of the
protein encoded thereby. The sequence of the human
protein was found to be as follows.

30

> (H₂N)-Ala-Lys-Tyr-Asn-Lys-Ile-Lys-Ser-Arg-Gly-
>   Ile-Lys-Ala-Asn-Ala-Phe-Lys-Lys-Leu-Asn-
>   Asn-Leu-Thr-Phe-Leu-Tyr-Leu-Asp-His-Asn-
>   Ala-Leu-Glu-Ser-Val-Pro-Leu-Asn-Leu-Pro-
>   Glu-Ser-Leu-Arg-Val-Ile-His-Leu-Gln-Phe-

35

-11-

Asn-Asn-Ile-Ala-Ser-Ile-Thr-Asp-Asp-Thr-
Phe-Cys-Lys-Ala-Asn-Asp-Thr-Ser-Tyr-Ile-
Arg-Asp-Arg-Ile-Glu-Glu-Ile-Arg-Leu-Glu-
Gly-Asn-Pro-Ile-Val-Leu-Gly-Lys-His-Pro-
5   Asn-Ser-Phe-Ile-Cys-Leu-Lys-Arg-Leu-Pro-
Ile-Gly-Ser-Tyr-Phe-(COOH),

A comparison of the human sequence with the
bovine sequence shows that there are seven differences at
10 positions 15, 20, 23, 54, 84, 85 and 105 of the sequence.
Accordingly, at least the residues at those positions may
be interchanged. It is possible, of course, that
sequences of other mammalian or avian species may exhibit
other differences.
15   In the course of obtaining the genes for the
mature bovine and human proteins it was discovered that
the genes each encode a precursor segment. Portions of
the precursor segments for the bovine and human proteins
are shown in Figure 11, parts A and B. Accordingly, it is
20 believed that the protein occurs as a prepropolypeptide
and is processed into the mature protein defined by the
sequences indicated above. Polypeptides comprising the
mature protein sequence and including a portion or all of
the precursor segments are intended to be within the scope
25 of the invention.
         The invention provides the protein in
substantially pure form in which it is essentially free of
other molecules with which it is associated in nature. In
this regard, the term "substantially pure" intends a
30 composition containing less than about 30% by weight
contaminating protein, preferably less than about 10%
contaminating protein, and most preferably less than about
5% by weight contaminating protein. The term
"substantially pure" is used relative to proteins with
35 which the protein is associated in nature and is not
intended to exclude compositions in which the protein is

-12-

admixed with nonproteinaceous carriers or vehicles, or
proteinaceous carriers or vehicles, provided other
protein(s) with which it is associated naturally are
absent. The invention also provides the protein in novel

5 partially glycosylated or totally deglycosylated forms
(both of which are referred to herein as
"deglycosylated").

            Based on the above amino acid sequences,
oligonucleotide probes which contain the codons for a por-

10 tion or all of the determined amino acid sequences are
prepared and used to screen DNA libraries for
substantially homologous genes that encode related
proteins. The homologous genes may be from other species
of mammals or animals (e.g., avians) or may represent

15 other members of a family of related genes. The basic
strategies for preparing oligonucleotide probes and DNA
libraries, as well as their screening by nucleic acid
hybridization, are well known to those of ordinary skill
in the art. See, e.g., DNA CLONING: VOLUME I (D.M. Glover

20 ed. 1985); NUCLEIC ACID HYBRIDIZATION (B.D. Hames and S.J.
Higgins eds. 1985); OLIGONUCLEOTIDE SYNTHESIS (M.J. Gate
ed. 1984); T. Maniatis, E.F. Frisch & J. Sambrook,
MOLECULAR CLONING: A LABORATORY MANUAL (1982).
            First, a DNA library is prepared. Since the

25 initially identified protein was bovine, it was logical to
probe a bovine library first, find full length clones and
use the full length bovine clones to probe libraries of
other mammalian species to identify the protein gene (and
thus the amino acid sequences) of other species. The

30 library can consist of a genomic DNA library. Bovine and
human genomic libraries are known in the art. See, e.g.,
Lawn et al., Cell (1978) 15:1157-1174. DNA libraries can
also be constructed of cDNA prepared from a poly-A RNA
(mRNA) fraction by reverse transcription. See, e.g., U.S.

35 Patent Nos. 4,446,235; 4,440,859; 4,433,140;

-13-

4,431,740; 4,370,417; 4,363,877. The mRNA is isolated
from an appropriate cell line or tissue that expresses the
factor. Libraries from cells involved in bone formation
(e.g., osteoblasts) or from osteotumors (e.g.,

5 osteosarcoma lines) are likely sources to probe for the
nucleic acids that encode the protein. cDNA (or genomic
DNA) is cloned into a vector suitable for construction of
a library. A preferred vector is a bacteriophage vector,
such as phage lambda. The construction of an appropriate

10 library is within the skill of the art.
        Once the library is constructed, oligonucleo-
tides to probe the library are prepared and used to
isolate the desired genes. The oligonucleotides are
synthesized by any appropriate method. The particular
nucleotide sequences selected are chosen so as to

15 correspond to the codons encoding the known amino acid
sequences of the protein. Since the genetic code is
redundant, it will often be necessary to synthesize
several oligonucleotides to cover all, or a reasonable
number, of the possible nucleotide sequences which encode

20 a particular region of a protein. Thus, it is generally
preferred in selecting a region upon which to base the
probes, that the region not contain amino acids whose
codons are highly degenerate. It may not be necessary,
however, to prepare probes containing codons that are rare

25 in the mammal from which the library was prepared. In
certain circumstances, one of skill in the art may find it
desirable to prepare probes that are fairly long, and/or
encompass regions of the amino acid sequence which would
have a high degree of degeneracy in corresponding nucleic

30 acid sequences, particularly if this lengthy and/or
degenerate region is highly characteristic of the protein.
Probes covering the complete gene, or a substantial part
of the genome, may also be appropriate, depending upon the
expected degree of homology. Such would be the case, for

35 example, if a cDNA of a bovine protein was used to screen

-14-

a human gene library for the corresponding human protein
gene. It may also be desirable to use two probes (or sets
of probes), each to different regions of the gene, in a
single hybridization experiment. Automated oligonucleo-
5   tide synthesis has made the preparation of large families
of probes relatively straightforward. While the exact
length of the probe employed is not critical, generally it
is recognized in the art that probes from about 14 to
about 20 base pairs are usually effective. Longer probes
10 of about 25 to about 60 base pairs are also used.

The selected oligonucleotide probes are labeled
with a marker, such as a radionucleotide or biotin using
standard procedures. The labeled set of probes is then
used in the screening step, which consists of allowing the
15 single-stranded probe to hybridize to isolated ssDNA from
the library, according to standard techniques. Either
stringent or permissive hybridization conditions could be
appropriate, depending upon several factors, such as the
length of the probe and whether the probe is derived from
20 the same species as the library, or an evolutionarily
close or distant species. The selection of the appropri-
ate conditions is within the skill of the art. See gener-
ally, NUCLEIC ACID HYBRIDIZATION, supra. The basic
requirement is that hybridization conditions be of suf-
25 ficient stringency so that selective hybridization occurs;
i.e., hybridization is due to a sufficient degree of
nucleic acid homology (e.g., at least about 75%), as op-
posed to nonspecific binding. Once a clone from the
screened library has been identified by positive
30 hybridization, it can be confirmed by restriction enzyme
analysis and DNA sequencing that the particular library
insert contains a gene for the protein.

Alternatively, a DNA coding sequence for a
protein can be prepared synthetically from overlapping
35 oligonucleotides whose sequence contains codons for the
amino acid sequence of the protein. Such oligonucleotides

-15-

are prepared by standard methods and assembled into a
complete coding sequence. See, e.g., Edge, Nature (1981)
292:756; Nambair et al., Science (1984) 223:1299; Jay et
al., J Biol Chem (1984) 259:6311.

5   Accordingly recombinant polynucleotides that
encode the polypeptides may be prepared and isolated by
one or more of the above described techniques. The term
"recombinant polynucleotide" as used herein denotes a
polynucleotide of genomic, cDNA, semisynthetic or
10 synthetic origin which, by virtue of its origin or
manipulation (1) is not associated with all or a portion
of the nucleic acid with which it is associated in nature
or in the form of a library, (2) is linked to a
polynucleotide to which it is not linked in nature or in a
15 library, or (3) is not found in nature or in a library.
           The DNA sequence coding for the protein can be
cloned in any suitable vector, identified, isolated, and
thereby maintained in a composition substantially free of
vectors that do not contain the coding sequence of the
20 protein (e.g., other library clones). Numerous cloning
vectors are known to those of skill in the art, and the
selection of an appropriate cloning vector is a matter of
choice. Examples of recombinant DNA vectors for cloning
and the host cells which they transform include
25 bacteriophage lambda (E. coli), pBR322 (E. coli),
pACYC177 (E. coli), pKT230 (gram-negative bacteria),
pGV1106 (gram-negative bacteria), pLAFR1 (gram-negative
bacteria), pME290 (non-E. coli gram-negative bacteria),
pHV14 (E. coli and Bacillus subtilis), pBD9 (Bacillus),
30 pIJ61 (Streptomyces), pUC6 (Streptomyces), actinophage C31
(Streptomyces), YIp5 (yeast), YCp19 (yeast), and bovine
papilloma virus (mammalian cells). See generally, DNA
CLONING: VOLUMES I & II, supra; MOLECULAR CLONING: A
LABORATORY MANUAL, supra.
35   In one embodiment of the present invention, the
coding sequence for gene encoding the protein is placed

-16-

under the control of a promoter, ribosome binding site
(for bacterial and eucaryotic expression), and optionally
an operator (collectively referred to herein as "control"
sequences), so that the DNA sequence encoding the protein
5   (referred to herein as the "coding" sequence) is
transcribed into RNA and the RNA translated into protein
in the host cell transformed by the vector. The coding
sequence may or may not contain a signal peptide or leader
sequence. The determination of the point at which the
10   precursor protein begins and the signal peptide ends is
easily determined from the N-terminal amino acid sequence
of the precursor protein. The protein can also be
expressed in the form of a fusion protein, wherein a
heterologous amino acid sequence is expressed at the
15   N-terminal. See, e.g., U.S. Patents Nos. 4,431,739;
4,425,437.
        The recombinant vector is constructed so that
the protein coding sequence is located in the vector with
the appropriate control sequences, the positioning and
20   orientation of the protein coding sequence with respect to
the control sequences being such that the coding sequence
is transcribed under the control of the control sequences
(i.e., by RNA polymerase which attaches to the DNA
molecule at the control sequences). The control sequences
25   may be ligated to the coding sequence prior to insertion
into a vector, such as the cloning vectors described
above. Alternatively, the coding sequence can be cloned
directly into an expression vector which already contains
the control sequence and an appropriate restriction site
30   downstream from control sequences. For expression of the
protein coding sequence in procaryotes and yeast, the
control sequences will be heterologous to the coding
sequence. If the selected host cell is a mammalian cell,
the control sequences can be heterologous or homologous to
35   the protein coding sequence, and the coding sequence can
be genomic DNA, cDNA or synthetic DNA. Either genomic or

-17-

cDNA coding sequence may be expressed in yeast. If
secretory expression in eukaryotic cells is necessary or
desirable, sequences such as the yeast alpha factor signal
sequence or other sequences that direct secretion are
5   included in the control sequence. If glycosylation
similar to the native molecule is desired, the gene may be
expressed in yeast or mammalian cells (COS, CHO, or CV-1
cells) using vectors and procedures known in the art.
A number of procaryotic expression vectors are known in
10   the art. See, e.g., U.S. Patent Nos. 4,440,859;
4,436,815; 4,431,740; 4,431,739; 4,428,941; 4,425,437;
4,418,149; 4,411,994; 4,366,246; 4,342,832. See also
British Patent Specifications GB 2,121,054; GB 2,008,123;
GB 2,007,675; and European Patent Specification 103,395.
15   Yeast expression vectors are known in the art. See, e.g.,
U.S. Patent Nos. 4,446,235; 4,443,539; 4,430,428. See
also European Patent Specifications 103,409; 100,561; and
96,491.
          Recombinant protein can be produced by growing
20   host cells transformed by the expression plasmid described
above under conditions whereby the protein is produced.
The protein is then isolated from the host cells and
purified. If the expression system secretes protein into
growth media, the protein can be purified directly from
25   cell-free media. If the recombinant protein is not
secreted, it is isolated from cell lysates. The selection
of the appropriate growth conditions and recovery methods
are within the skill of the art or are apparent from the
recovery methods used to isolate the native proteins. The
30   recombinant protein may be recovered by affinity
chromatography using the antibodies produced in accordance
with the invention. Recombinant protein may be
unglycosylated or have a different glycosylation pattern
than the native molecule depending upon the host that is
35   used to produce it. The proteins are useful for making

-18-

antibodies that recognize sequential epitopes of the
protein, and are useful as bone markers.

Either native, deglycosylated, or synthetic
(recombinant) protein can be used to produce antibodies,
5 both polyclonal and monoclonal. The term "antibody" is
intended to include whole Ig of any isotype or species as
well as antigen binding fragments, chimeric constructs and
single chain antibodies. If polyclonal antibodies are
desired, purified protein is used to immunize a selected
10 mammal (e.g., mouse, rabbit, goat, horse, etc.) and serum
from the immunized animal later collected and treated
according to known procedures. Compositions containing
polyclonal antibodies to a variety of antigens in addition
to the protein can be made substantially free of anti-
15 bodies which do not bind specifically to the protein
bodies by passing the composition through a column to
which protein has been bound. After washing, polyclonal
antibodies to the protein are eluted from the column.
Monoclonal anti-protein antibodies can also be readily
20 produced by one skilled in the art. The general methodol-
ogy for making monoclonal antibodies by hybridomas is well
known. Immortal, antibody-producing cell lines can also
be created by techniques other than fusion, such as direct
transformation of B lymphocytes with oncogenic DNA, or
25 transfection with Epstein-Barr virus. See, e.g.,
Schreier, M., et al., HYBRIDOMA TECHNIQUES (1980);
Hammerling et al., MONOCLONAL ANTIBODIES AND T-CELL
HYBRIDOMAS (1981); Kennett et al., MONOCLONAL ANTIBODIES
(1980).

30   By employing the bone-specific protein (native,
deglycosylated or synthetic) as an antigen in the im-
munization of the source of the B-cells immortalized for
the production of monoclonal antibodies, a panel of
monoclonal antibodies recognizing epitopes at different
35 sites on the protein molecule can be obtained. Antibodies
which recognize an epitope in the binding region of the

protein can be readily identified in competition assays between antibodies and protein. Antibodies which recognize a site on the protein are useful, for example, in the purification of the protein from cell lysates or
5 fermentation media, in characterization of the protein and in identifying immunologically related proteins. Such immunologically related proteins (i.e., that exhibit common epitopes with the protein) are another aspect of the invention. In general, as is known in the art, the
10 anti-protein antibody is fixed (immobilized) to a solid support, such as a column or latex beads, contacted with a solution containing the protein, and separated from the solution. The protein, bound to the immobilized antibodies, is then eluted. Antibodies to the protein may
15 be used to identify osteoblasts and osteocytes by conventional immunoassay procedures. Such identification may be used to follow bone and/or cartilege turnover.

20 Examples

        The following is intended to further illustrate processes for preparing the proteins of the invention and their use in preparing antibodies. These examples are not intended to limit the invention in any manner.
25

A. Preparation of Demineralized Bone

        Bovine metatarsal bone was obtained fresh from the slaughterhouse and transported on ice. Bones were cleaned of all periosteum and marrow with high pressure
30 water, crushed into fragments using a liquid-nitrogen-cooled grinder and pulverized into powder using a liquid-nitrogen-cooled mill. The pulverized bone was washed four times for 20 minutes in $4^{O}C$ deionized water (8 liters/kg). The bone was then washed overnight with the same volume of
35 deionized water at $4^{O}C$. The bone powder was demineralized for 5 hr in 0.5 N HCl (21 liter/kg) at $4^{O}C$. The acid was

-20-

decanted, and the demineralized bone powder was washed
several times with 4°C deionized water until the wash
reached a pH>3. The excess water was removed on a suction
filter.

5

B. Extraction of Noncollagenous Proteins
        Demineralized bone powder was extracted with 4M
guanidine-HCl, 10 mM EDTA pH 6.8 (2 liters/kg bone powder)
for 16 hr at 4°C. The suspension was suction-filtered to
10 recover the guanidine-HCl-soluble fraction and
concentrated at least 5-fold by ultrafiltration using a
10,000 dalton cut-off membrane (S10Y10 Amicon spiral
cartridge).

15 C. Gel Filtration
        The extract from ¶B, redissolved in 4M
guanidine-HCl, was fractionated on a Sephacryl S-200
column equilibrated in 4M guanidine-HCl, 0.02% sodium
azide, 10 mM EDTA, pH 6.8. Fractions were assayed by
20 their absorbance at 280 nm and the fractions were combined
as shown in Figure 2. The fraction indicated by <----> in
Figure 2 constitutes a low molecular weight (LMW, 10,000-
30,000 daltons) protein fraction possessing the greatest
activity. This fraction was pooled and dialyzed against 6
25 changes of 180 volumes of deionized water and lyophilized.
All operations except lyophilization and dialysis (4°C)
were conducted at room temperature.

D. Ion Exchange Chromatography
30   The pooled fraction from ¶C was dissolved in 6M
urea, 10 mM NaCl, 1 mM NEM, 50 mM sodium acetate, pH 4.8
and centrifuged at 10,000 rpm for 5 min. The supernatant
was fractionated on a CM52 (a commercially available CMC)
column equilibrated in the same buffer. Bound proteins
35 were eluted from the column using a 10 mM to 400 mM NaCl
gradient in the same buffer, and a total volume of 350 ml

-21-

at a flow rate of 27 ml/hr. Proteins eluted with 10-150
mM NaCl (the <----> of Figure 3) were collected and
dialyzed against 6 changes of 110 volumes of deionized
water for 4 days and lyophilized. All of the foregoing
5 operations were conducted at room temperature except
dialysis (4°C).

E.   ConA Chromatography
          The fraction obtained in step D above was
10 enriched by affinity chromatography using concanavalin A
(ConA)-Sepharose 4B (Pharmacia). In order to minimize
leaching of ConA from the column during chromatography,
the resin was cross-linked with.glutaraldehyde essentially
as described by K.P. Campbell, D.H. MacLennan, J Biol Chem
15 (1981) 256:4626. Briefly, resin was pelleted (500 x g,
5 min) and washed twice with 4 volumes of 250 mM NaHCO$_3$,
pH 8.8. The resin was then equilibrated in the same
buffer for 6-8 hrs at 4°C. After pelleting, the resin was
cross-linked by the addition of 4 volumes of 250 mM
20 NaHCO$_3$, pH 8.8, 250 mM methyl-alpha-D-mannopyranoside
(alpha-MM), 0.03% glutaraldehyde with gentle mixing for
1 hr at room temperature. The reaction was quenched by
washing the resin twice in 1M Tris-HCl, pH 7.8. The resin
was stored in the same buffer containing 0.01% Thimersol
25 at 4°C until use.
          Samples for ConA chromatography were solubilized
in 1% deoxycholate at pH 8.0. Any small amount of
precipitate was removed by centrifugation 12,000 x g, 5
minutes.
30   Prior to chromatography, cross-linked resin was
first equilibrated with >5 column volumes of 50 mM Tris,
pH 8.0 followed by >5 column volumes of 1% sodium
deoxycholate. Samples were loaded and nonbound fractions
collected by washing with 1% DOC. Elution was monitored
35 by OD$_{280}$. Bound material was eluted with 0.5M alpha-MM in
1% DOC as shown in Figure 4.

-22-

F.   Chromatography on Heparin-Sepharose

        The bound fraction eluted from the ConA column
was reequilibrated by chromatography on a GH-25 column
5   (Pharmacia) equilibrated in 6M urea, 0.1M NaCl, 50 mM
Tris-HCl pH 7.2 heparin-sepharose buffer. Approximately
80 mg (1 mg/ml) were loaded on a 25 ml large heparin
sepharose column (Pharmacia). The column was washed of
all unbound material. Then bound proteins were eluted
10 with the same equilibrating buffer but containing 0.5M
NaCl as shown in Figure 5. About 5-8 mg of heparin-
sepharose bound proteins were recovered.

G.   Chromatography on C18-RP-HPLC
15   The pH of the heparin-bound fraction was lowered
below 5 by adding TFA. Final purification of the heparin-
bound fraction was achieved using reversed phase HPLC.
The columns used were a Vydac TP-RP18 4.6 mm x 25 cm and
1.0 x 25 cm. Solvent A was 0.1% aqueous trifluoroacetic
20 acid (TFA) and B 90% acetonitrile in A. Bound proteins
were eluted from the column with a 32-62% B solvent gradi-
ent at a rate of 1%/min. The protein composition eluted
between 47-50% solvent B as shown in Figure 6. 140-200 ug
protein were recovered. Amino acid composition and amino
25 acid sequences of the protein were determined using
standard procedures and are described above and shown in
Figure 7.

H.   Deglycosylation
30   Glycopeptidase F cleaves N-linked oligo-
saccharides at the innermost N-acetylglucosamine residue.
High mannose, hybrid and complex oligosaccharides are
susceptible to the enzyme. Protein was iodinated by the
chloramine-T method. Labeled protein was digested for 12-
35 15 hours with 6.7 units/ml glycopeptidase F (Boehringer
Mannheim) in 0.1M Tris-HCl, pH 7.4, 10 mM EDTA, at 37°C.

-23-

         Both the glycosylated and deglycosylated forms
were analyzed by sodium dodecyl sulfate/15% polyacrylamide
slab gels prepared according to standard methods. Figure
8 is a photograph of the autoradiograph.

5

I. Isolation of Bovine Protein Gene
         The protein is designated "OIF" in the drawings.
         The following four 20-mer oligonucleotide probes
were synthesized using a Biosearch 8600 DNA synthesizer.
10 The sequences of these probes were derived from the amino
acid sequence of the bovine protein that was isolated from
bone.

         GCNAARTAYAAYAARATQAA
15   TTYCGNTTRTGNAARTTYTT
         CTRCTRTGNAARACRTTYCG
         CTYCCNTTRGGNCANTAWGA

where A is adenine, C is cytosine, G is guanine, T is
20 thymine, N is A, C, G or T, Q is A, C, or T, R is A or G,
W is A, G or T and Y is C or T.
         These probes were used to analyze a lambda
bacteriophage "library" containing DNA fragments from
bovine liver. The lambda phage vector, EMBL3 (Frischauf,
25 A.M., et al., J Mol Biol (1983) 170:827) was purchased
(Stratagene, 1190 North Torrey Pines Road, La Jolla, CA
92037) and used as described. Bovine liver was collected
at a slaughterhouse and quickly frozen in liquid nitrogen.
The frozen tissue was pulverized and lysed with sarkosyl
30 NL-97A and proteinase K. Cellular DNA was purified by
CsCl density gradient centrifugation, treated with Sau3A,
and fractionated by sucrose gradient centrifugation after
phenol-chloroform extraction.
         The DNA was concentrated to 1 mg/ml and 1 ug was
35 mixed with 1 ug EMBL3 DNA. The mixture was treated with
DNA ligase as described by the supplier and packaged via

-24-

the "gigapack kit" (Stratagene) to make a library stock.
Approximately 10,000 viable phage were plated on each of
60 plates (150 mm) (see Molecular Cloning: A Laboratory
Manual, Maniatis, Fritsch and Sambrook, Cold Spring Harbor
5   Laboratory, Cold Spring Harbor, New York (1983)).
          Phage plaques were transferred to nitrocellulose
filters (4 replicates per plate). Absorbed DNA was de-
natured by treatment with 0.5M NaOH, 1.5M NaCl. Filters
were neutralized in 0.5M Tris, pH 8.0, 1.5M NaCl, washed
10   with 2XSSC, air dried, and baked for 2 hr at 80°C under
vacuum.
          Phage containing the gene were identified by
hybridizing ^{32}P-labeled oligonucleotides with filters
containing DNA transferred from phage plaques. Plaques
15   hybridizing in duplicate to at least two of the four
oligonucleotides were purified and further characterized.
From three experiments representing a total of 6 x 10^5
plaques (2-4 bovine genome equivalents) only two plaques
were shown on final analysis to contain sequences compat-
20   ible with the protein sequence. These two phage (bOIF21
and bOIF39) were shown by restriction site mapping and
Southern analysis to contain identical sequences in the
bOIF region although the extent of bovine DNA in the two
was different.
25   Sequencing of these clones revealed the gene
structure shown in Figure 9 and the DNA sequence of Figure
11. As shown, the mature bovine sequence is encoded by
three exons, designated 1, 2, and 3 in Figure 9. Exon 1
encodes the first 17 amino acids, Exon 2 amino acids 18
30   through 49, and Exon 3 the remaining residues. The pre-
cursor segment of the prepropolypeptide is encoded by a
portion of Exon 1 and separate exons upstream from Exon 1.
          To confirm the splicing of the exons, a bovine
cancellous bone mRNA library was prepared using the
35   polymerase chain reaction (PCR) with primers selected from
within Exons 1 and 3. A clone containing the bOIF cDNA

-25-

sequence was isolated and the nucleotide sequence
determined. The resulting sequence confirmed the exon
splicing contemplated by Figures 9 and 11.

5 J.   Isolation of Human Protein Gene
            Human fetal liver DNA was isolated, treated with
    Sau3A as was described for bovine DNA (above). Phage
    plaques (2.5 x 10^5) form the EMBL3 human liver DNA library
    (generated as described above) and a similar number of
10 plaques from a human liver DNA library produced in lambda
    phage Charon 4a (Lawn, R.M., et al., Cell (1978) 15:1157)
    were probed with a radioactively-labeled (EcoRI) DNA frag-
    ment (Fragment 1, Figure 9) containing the first exon of
    the bovine gene. All positive appearing plaques (13 from
15 the EMBL3 library and 6 from the Charon 4a) were isolated
    and reprobed with the radioactive exon 1 DNA fragment as
    well as a second DNA fragment containing exon 3 (Fragment
    2, Figure 9) of the bovine gene. Only one phage from the
    EMBL3 library (phage 41) hybridized to both DNA fragments
20 and a second phage (phage 28) from the gene library
    hybridized with only the exon 1 fragment upon rescreening.
    All other phage did not hybridize with the radioactive
    probes upon rescreening, indicating that their original
    identification was either an artifact of hybridization or
25 that the human protein cognate DNA was lost upon
    replication during isolation of the phage.
            Phage 41 and 28 were subjected to restriction
    site mapping and Southern blot analysis to locate the
    sequences for the human probes and the DNA sequence of
30 these regions was determined.
            Sequencing of these clones showed that the human
    gene structure paralleled that shown in Figure 9, that is
    the human gene for the mature protein comprises three
    exons of identical length to the bovine exons. A restric-
35 tion map of the human gene region is shown in Figure 10.
    It was further determined that the human protein, as the

-26-

bovine, occurs as a prepropolypeptide with the precursor
segment being encoded by a portion of Exon 1 and upstream
exons. The human genomic DNA sequence is shown in Figure
11.

5

K.   Construction of Mammalian Expression Vector Containing
     Human Protein Gene
          A HindIII fragment of the human protein gene
     (site 15422 to site 25814 in Figure 10) was cloned into a
10   HindIII site of plasmid pSC614 (see Figure 12) in
     transcriptional alignment with the SV40 promoter or in the
     opposite orientation to yield plasmids phOIF17 and phOIF18
     respectively. Figure 12 is a schematic diagram of plasmid
     phOIF18.

15

L.   Transfection of COS-7 Cells with phOIF18
          COS-7 cells were transfected with phOIF18 (see
     Felgner, P.L., et al., Proc Natl Acad Sci (USA) (1987)
     84:7413). After transfection the cells were allowed to
20   grow in medium with or without serum (5%). In these
     experiments only the cells grown in serum-containing media
     synthesized the protein.

M.   Production and Testing of Antibodies to the Protein
25   M.1.   Production of Polyclonal Antibodies
          Polyclonal antibodies to (1) a synthetic 30-mer
     polypeptide having a sequence corresponding to the amino
     acids 1-30 of Figure 7 except for a Leu ----> Asn
     substitution at position 25 and (2) the native protein
30   purified from bone as described above were prepared and
     characterized as follows.
          Antiserum to the 1-30-mer was raised in a rabbit
     by injecting the rabbit with 500 ug of the polypeptide in
     complete Freund's adjuvant (CFA), followed by boosts of
35   500 ug of the polypeptide in incomplete Freund's adjuvant
     (ICFA) at approximately three week intervals. The

-27-

antiserum was obtained after the fourth boost and had a
titer as measured by ELISA of >1:10,000. Rabbit antiserum
to the native protein was raised similarly using an
initial injection of 50 ug protein in CFA followed by
5   boosts of 50 ug protein in ICFA. This antiserum had a
titer of >1:10,000 by ELISA.

        The antiserum to the 1-30-mer was tested in
Western blots on the purified native protein,
deglycosylated native protein, and on crude native protein
10 (Con-A bound material), all fixed post-blotting with 0.2%
glutaraldehyde. The antiserum detected the purified
native protein at >1 ug and also recognized the
deglycosylated protein and the crude protein. The
antiserum to the native protein recognized the native
15 protein at >100 ng in Western blots.

        M.2. Production of Monoclonal Antibodies
        Murine monoclonal antibodies to the purified
native protein were prepared as follows. From two fusions
20 25 positive wells were identified by immunoprecipitation.
A group of female Balb/c mice was injected
intraperitoneally (IP) with 10-20 ug of purified native
protein in CFA. The animals were boosted with 10-20 ug of
protein in ICFA. Following the third boost, the mice were
25 bled and serum antibody titers against the protein checked
by ELISA. Two animals were found to have titers of
>1:40,000. They were given a final intravenous (IV)
injection of 20 ug protein four days prior to the fusion.
        Fusion to the SP2/0 myeloma (GM3659 B, NIGMS
30 Human Genetic Mutant Cell Repository, Camden, NJ) was
performed essentially according to the protocol of Oi and
Herzenberg, "Immunoglobulin-producing Hybrid Cell Lines"
in Selected Methods in Cellular Immunology, Mishell and
Shiigi, eds., W.H. Freeman and Co., San Francisco, pp.
35 357-362, (1980). Spleen cells from the animals were mixed
with SP2/0 at a ratio of 5:1. 50% polyethylene glycol

-28-

1500 (Boehringer-Mannheim Biochemicals, Indianapolis, IN) was used as the fusagen. Cells were plated at 10^6 cells/ well along with resident peritoneal cells at 4×10^3 cells/well in DMEM with high glucose (4.5 g/l) sup-
5 plemented with 20% FCS (Hyclone Laboratories, Logan, UT), 2 mM L-glutamine, 2 mM sodium pyruvate, nonessential amino acids, penicillin and streptomycin. In this procedure, aminopterin was replaced by azaserine (Sigma) according to the procedure by Larrick et al., Proc Natl Acad Sci (USA)
10 (1983) 80:6376, and added along with thymidine and hypoxanthine on day 1 after the fusion.

From two fusions 25 positive wells were identi- fied by immunoprecipitation of ^{125}I-labeled protein.

All 25 were also positive in an ELISA against
15 the protein. In addition, several other wells were positive by ELISA but negative by immunoprecipitation. The supernatant from one uncloned well (3B2.17, previously designated FO13-3B2) was particularly positive and was used in a Western blot. In this testing synthetic
20 peptides corresponding to amino acid segments 1-30, 62-95 and 76-105 of the protein sequence were made and 1-2 ug of each was applied to separate lanes in the gel. Blots were probed with 50-100 ug/ml of purified antibody. This anti- body recognized ≥300 ng protein as well as deglycosylated
25 protein. The antibody also picked up the protein in a crude fraction (total Con-A bound) and was found to recognize the C-terminal peptide (76-105) but not the N- terminal peptide (1-30). Another clone, designated 2C11.6, was found to recognize the internal 62-95 segment.
30 Clones 3B2.17 and 2C11.6 were subcloned by limiting dilu- tion and were found to be stable and to be IgG isotype. These clones have been deposited in the American Type Culture Collection (ATCC) on 5 April 1989 under the provi- sions of the Budapest Treaty. Their ATCC designations
35 are, respectively, HB10099 (3B2.17) and HB10098 (2C11.6).

-29-

M.3. Immunostaining With Antibodies

Rat fetuses (19 days old) and 3 day old rats
were used for tissue sections. Tissue was fixed in 10%
formalin and 5u sections were prepared in paraffin. The
5 sections were treated with xylene (deparaffinized) and
washed with Tris-buffered saline (TBS), pH 7.6, three
times. The washed sections were then contacted with a
mixture of TBS, 0.05% Tween, 0.5% bovine serum albumin
(BSA), 10% normal mouse serum (NMS), for 1 hr at 25°C or
10 overnight at 4°C. Monoclonal antibody 3B2 (see M.2) at 5
ug/ml in TBS/Tween/BSA/NMS was added and the sections were
incubated for 1 hr at room temperature. The sections were
then washed three times in TBS/Tween and contacted with
biotinylated goat anti-mouse antibody for 10 min at room
15 temperature. The sections were then washed again three
times with TBS/Tween and contacted with streptavidin-
horseradish peroxidase for 10 min at room temperature.
Thereafter, the sections were washed a final three times
with TBS/Tween and contacted with substrate. After
20 substrat. treatment, the sections were washed in water,
counterstained with Mayer's hematoxylin, washed again in
water, dehydrated in 100% ethanol, and mounted.

The most intenst staining occurred in
hypertrophic calcifying cartilage in the growth plate.
25 There was also a clear staining pattern in osteoblasts and
osteocytes with the darkest in more mature cells. No
detectable staining was observed in soft tissue. A faint
pattern was seen in bone itself.

Modifications of the above-described modes of
30 carrying out the invention that are obvious to those of
skill in the arts relevant to the invention are intended
to be within the scope of the following claims.

35

-30-

Claims

1. A substantially pure polypeptide having the
following amino acid sequence:

(H_2N)-Ala-Lys-Tyr-Asn-Lys-Ile-Lys-Ser-Arg-Gly-
      Ile-Lys-Ala-Asn-R^1-Phe-Lys-Lys-Leu-R^2-
      Asn-Leu-R^3-Phe-Leu-Tyr-Leu-Asp-His-Asn-
      Ala-Leu-Glu-Ser-Val-Pro-Leu-Asn-Leu-Pro-
      Glu-Ser-Leu-Arg-Val-Ile-His-Leu-Gln-Phe-
      Asn-Asn-Ile-R^4-Ser-Ile-Thr-Asp-Asp-Thr-
      Phe-Cys-Lys-Ala-Asn-Asp-Thr-Ser-Tyr-Ile-
      Arg-Asp-Arg-Ile-Glu-Glu-Ile-Arg-Leu-Glu-
      Gly-Asn-Pro-R^5-R^6-Leu-Gly-Lys-His-Pro-
      Asn-Ser-Phe-Ile-Cys-Leu-Lys-Arg-Leu-Pro-
      Ile-Gly-Ser-Tyr-R^7-(COOH),

where R^1 is Ala or Thr, R^2 is Asn or His, R^3 is Thr or
Ser, R^4 is Ala or Thr, R^5 is Ile or Val, R^6 is Val or Ile
and R^7 is Phe or Ile, and substantially pure polypeptides
that are and substantially homologous thereto or are
immunologically related thereto.

2. The polypeptide of claim 1 wherein R^1 is
Ala, R^2 is Asn, R^3 is Thr, R^4 is Ala, R^5 is Ile, R^6 is
Val, and R^7 is Phe.

3. The polypeptide of claim 1 wherein R^1 is
Thr, R^2 is His, R^3 is Ser, R^4 is Thr, R^5 is Val, R^6 is
Ile, and R^7 is Ile.

4. A substantially pure prepropolypeptide
comprising
      (a) the polypeptide of claim 1 and
      (b) at least a portion of a precursor segment
having the sequence shown in Figure 11A or Figure 11B,
      or a sequence substantially homologous thereto.

    5. A substantially pure prepropolypeptide
comprising
        (a)   the polypeptide of claim 2 and
 5   (b)   at least a portion of a precursor segment
having the human sequence shown in Figure 11, Parts A and
B,
        or a sequence substantially homologous thereto.

10   6. A substantially pure prepropolypeptide
comprising
        (a)   the polypeptide of claim 3 and
        (b)   at least a portion of a precursor segment
having the bovine sequence shown in Figure 11, Part B,
15   or a sequence substantially homologous thereto.

    7. A polypeptide having t e following amino
acid sequence:

20   (H_2N)-Ala-Lys-Tyr-Asn-Lys-Ile-Lys-Ser-Arg-Gly-
        Ile-Lys-Ala-Asn-R^1-Phe-Lys-Lys-Leu-R^2-
        Asn-Leu-R^3-Phe-Leu-Tyr-Leu-Asp-His-Asn-
        Ala-Leu-Glu-Ser-Val-Pro-Leu-Asn-Leu-Pro-
        Glu-Ser-Leu-Arg-Val-Ile-His-Leu-Gln-Phe-
25   Asn-Asn-Ile-R^4-Ser-Ile-Thr-Asp-Asp-Thr-
        Phe-Cys-Lys-Ala-Asn-Asp-Thr-Ser-Tyr-Ile-
        Arg-Asp-Arg-Ile-Glu-Glu-Ile-Arg-Leu-Glu-
        Gly-Asn-Pro-R^5-R^6-Leu-Gly-Lys-His-Pro-
        Asn-Ser-Phe-Ile-Cys-Leu-Lys-Arg-Leu-Pro-
30   Ile-Gly-Ser-Tyr-R^7-(COOH),

where R^1 is Ala or Thr, R^2 is Asn or His, R^3 is Thr or
Ser, R^4 is Ala or Thr, R^5 is Ile or Val, R^6 is Val or Ile
and R^7 is Phe or Ile, wherein said polypeptide is
35 deglycosylated relative to a native polypeptide having
said sequence and deglycosylated polypeptides that are
substantially homologous thereto or are immunologically
related thereto.

-32-

8.   Antibody that binds to a polypeptide of claim 1, 2, 3, or 7.

5   9.   A recombinant polynucleotide encoding a polypeptide of claim 1, 2 or 3.

10.   A recombinant polynucleotide encoding a prepropolypeptide of claim 4, 5, or 6.

10

11.   A recombinant vector containing a recombinant polynucleotide of claim 9 and capable of directing the expression of the polypeptide encoded thereby.

15

12.   A recombinant vector containing a recombinant polynucleotide of claim 10 and capable of directing the expression of the prepropolypeptide encoded thereby.

20

13.   A recombinant host cell or microorganism containing the recombinant vector of claim 11 and capable of permitting expression of said polypeptide.

25   14.   A recombinant host cell or microorganism containing the recombinant vector of claim 12 and capable of permitting expression of said polypeptide.

30

35

DEMINERALIZED BONE POWDER (0.5 N HCL)

↓

4 M GUANIDINE HYDROCHLORIDE EXTRACT

↓

SEPHACRYLS – 200 (20 – 36KD MW)

↓

CM – CELLULOSE (10 – 150 mM NaCL)

↓

CONCANAVALIN – A (0.5 M METHYL ∝ – D – MANNOPYRANOSIDE)

↓

HEPARIN – SEPHAROSE (0.1 – 0.5 M NaCL)

↓

C 18 – REVERSED PHASE – HPLC (42 – 45% ACETONOTRILE)

FIG. 1
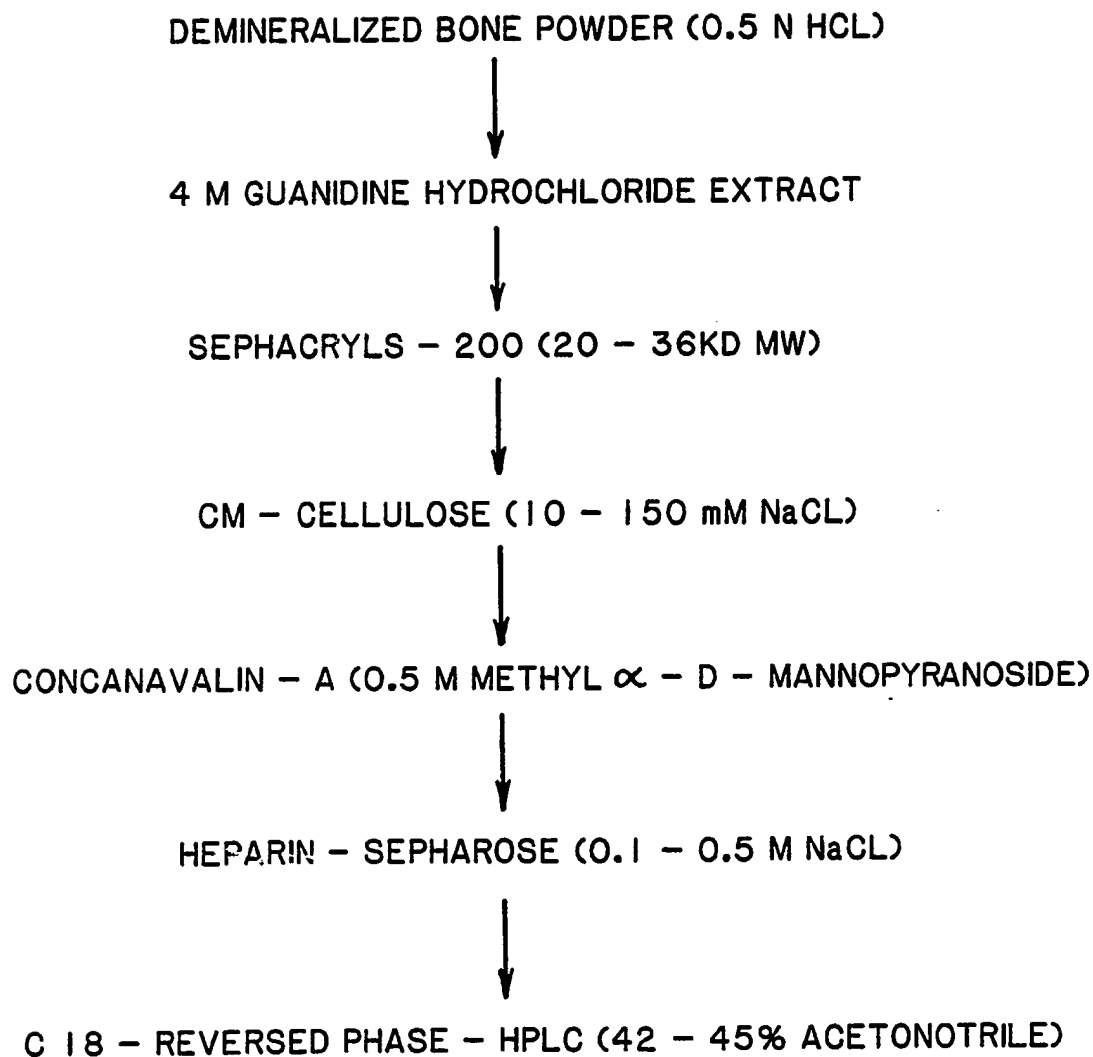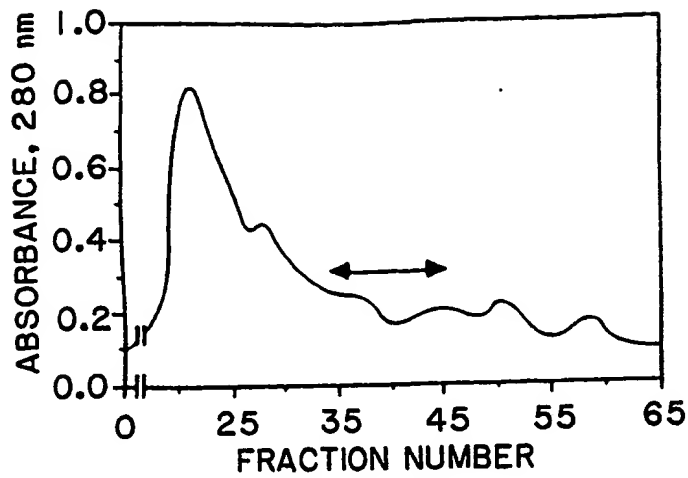
FIG. 2

FIG. 3

FIG. 4

FIG. 5

FIG. 6
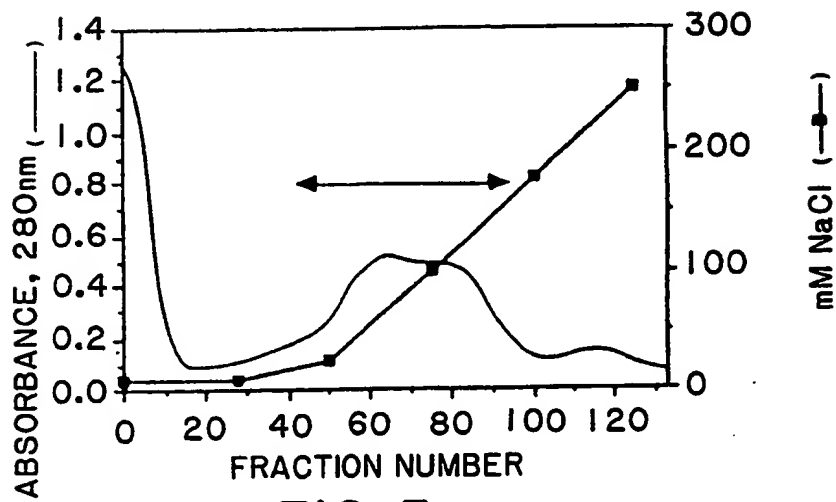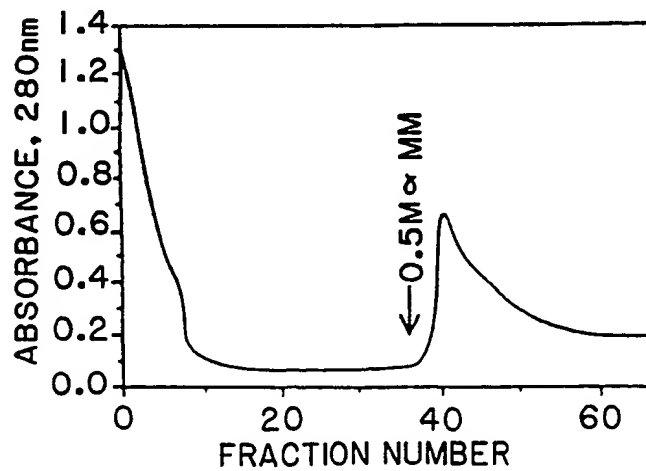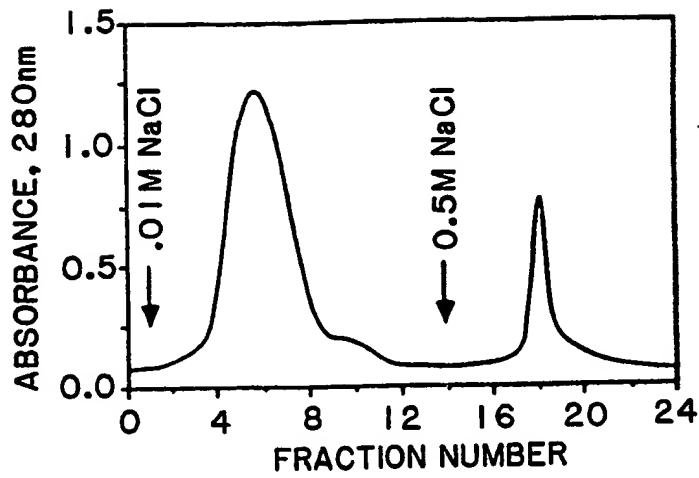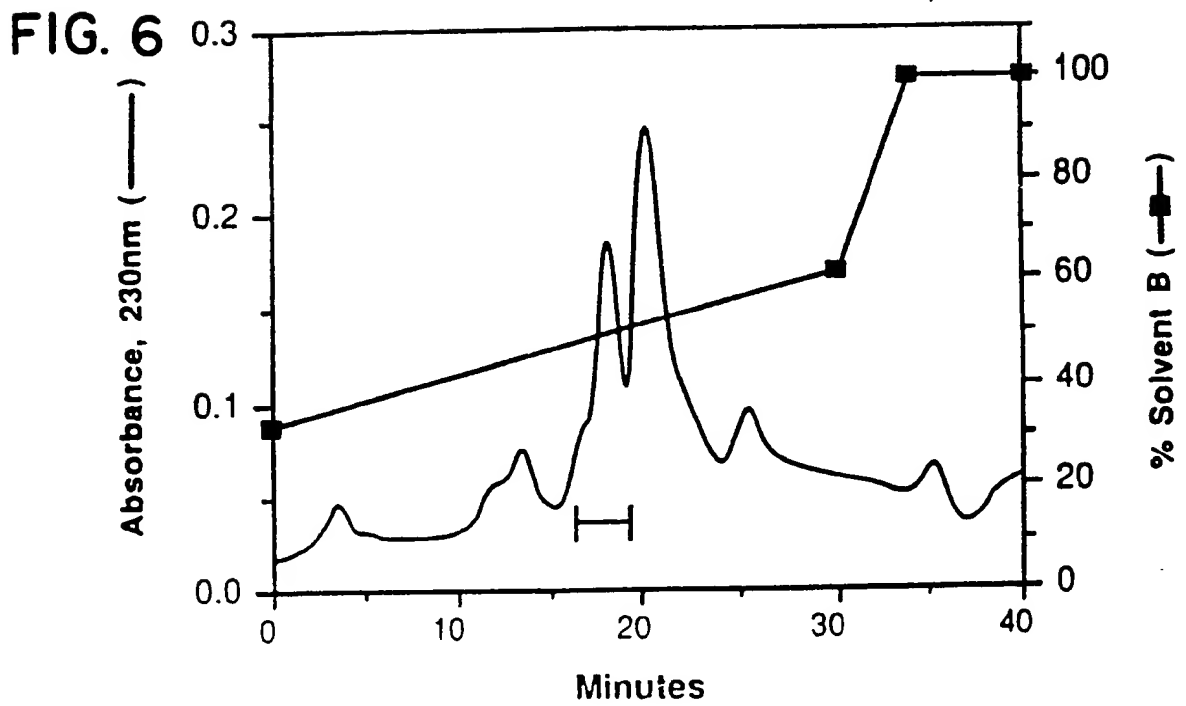
FIG. 7-1

AA #	N-	LYS-C	GLU-C	LYS-C	GLU-C	LYS-C
1	ALA					
2	LYS					
3	TYR					
4	ASN					
5	LYS					
6	ILE					
7	LYS					
8	SER					
9	ARG					
10	GLY					
11	ILE					
12	LYS					
13	ALA					
14	ASN					
15	THR					
16	PHE					
17	LYS					
18	LYS	(LYS)				
19	LEU	LEU				
20	HIS	HIS				
21	ASN	ASN				
22	LEU	LEU				
23	SER	SER				
24	PHE	PHE				
25	LEU	LEU				
26	TYR	TYR				
27	LEU	LEU				
28	ASP	ASP				
29	HIS	HIS				
30	ASN	ASN				
31	ALA	ALA				
32	LEU	LEU				
33	GLU	GLU				
34		SER				
35		VAL				
36		PRO				
37		LEU				
38		ASN				
39		LEU				
40		PRO				
41		GLU				
42		SER	SER			
43		LEU	LEU			
44			ARG			
45			VAL			
46			ILE			
47			HIS			
48			LEU			
49			GLN			
50			PHE			
51			ASN			
52			ASN			

FIG. 7-2

53			ILE		
54			THR		
55			SER		
56			ILE		
57			THR		
58			ASP		
59			ASP		
60			THR		
61			PHE		
62			CYS		
63			LYS		
64			ALA	ALA	
65				ASN-CHO	
66				ASP	
67				THR	
68				SER	
69				TYR	
70				ILE	
71				ARG	
72				ASP	
73				ARG	
74				ILE	
75				GLU	
76				GLU	
77				ILE	
78				ARG	
79				LEU	
80				GLU	
81				GLY	GLY
82				ASN	ASN
83				PRO	PRO
84				VAL	VAL
85				ILE	ILE
86				LEU	LEU
87				GLY	GLY
88				LYS	LYS
89					HIS
90					PRO
91					ASN
92					SER
93					PHE
94					ILE
95					CYS
96					LEU
97					LYS
98					ARG	ARG
99					LEU	LEU
100					PRO	PRO
101					ILE	ILE
102					GLY	GLY
103					SER	SER
104					TYR	TYR
105						ILE-COOH

FIG. 8

```
ENZ  -  +     -  +
      115B    115D
```

45K —

30K —

20K —

14K —

FIG. 9

FIG. 10

HIND3 15422
SPE1 17072
ECOR5 17403
XBA1 18572
NHE1 19379
BSSH2 20179
XBA1 20879

NHE1 1900
NHE1 4554
HIND3 4956
SPE1 5043
ECOR5 5815
HIND3 7515

HIND3 25814
XBA1 25870
XBA1 28343
HIND3 30512
HIND3 32459

N-terminal Exon->   Exon-1->   Exon2->
                              Exon1->
                                 Exon3->

HIND3 15408

dhfr

ORI
Amp

Exon-1

CAM

phOIF18
18848 bp
OIF

SV40ori

Exon3
Exon2

HIND3 4973

Exon1

Human DNA
Containing
OIF Gene
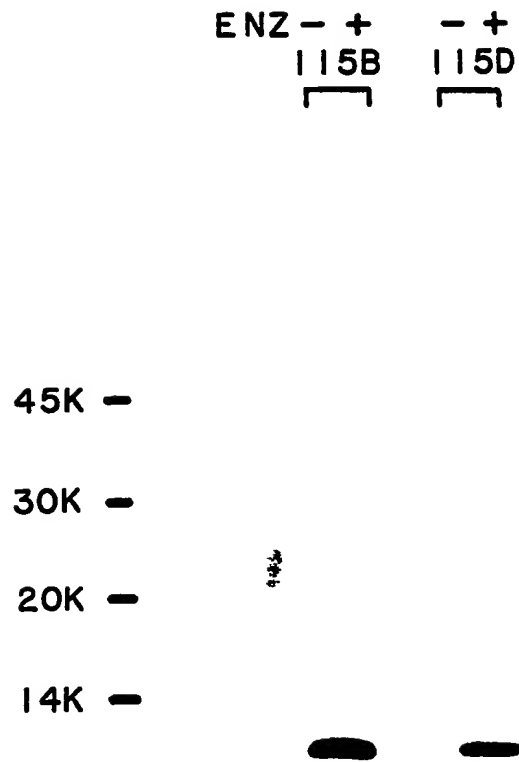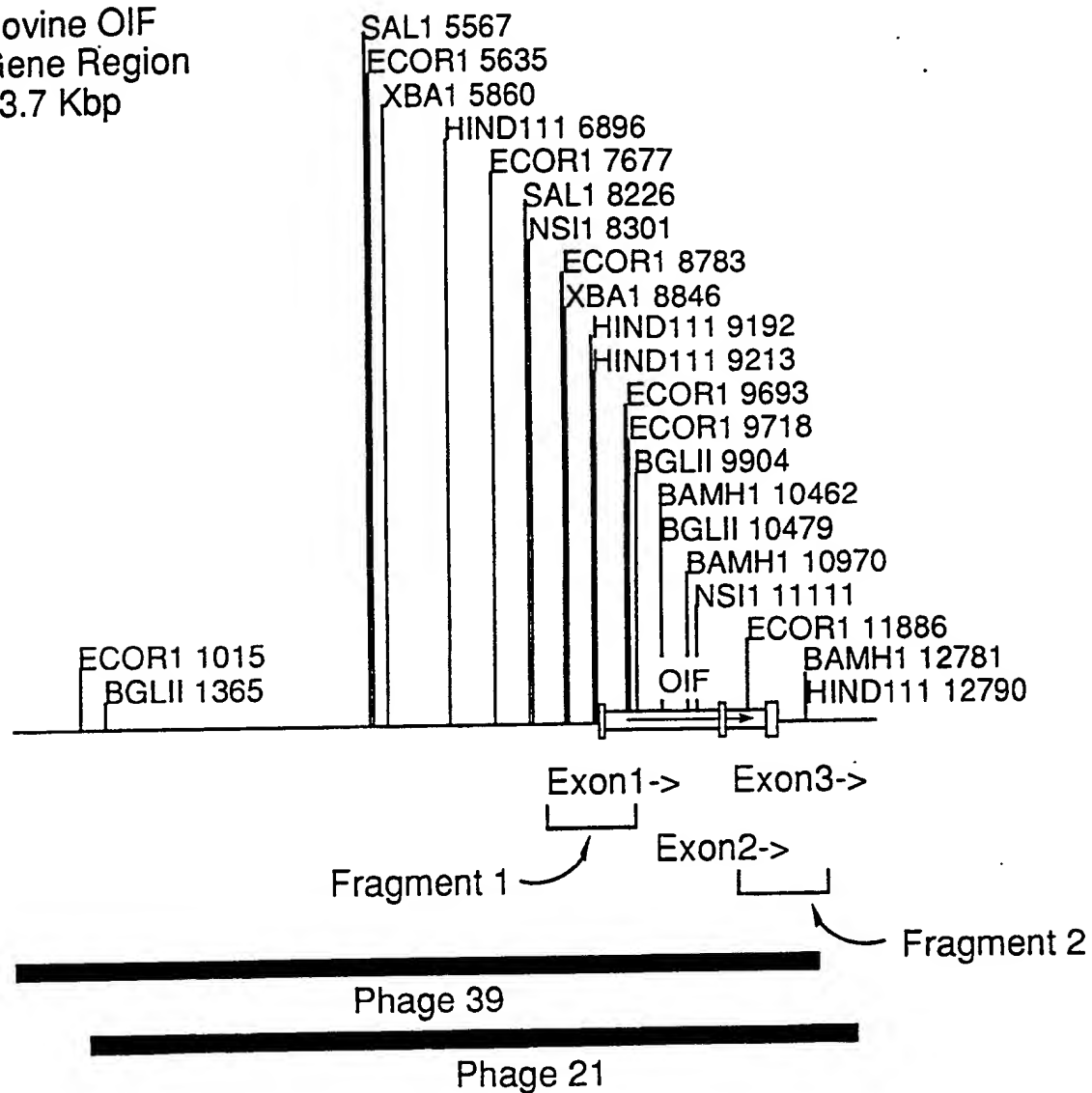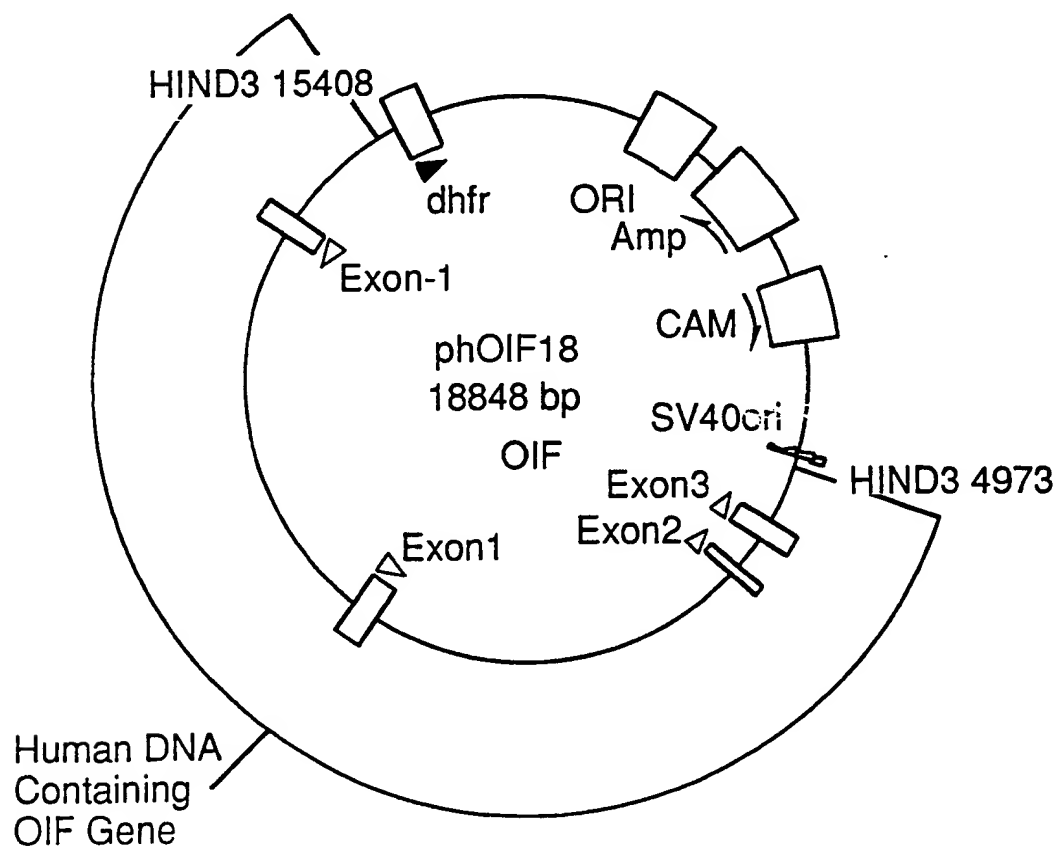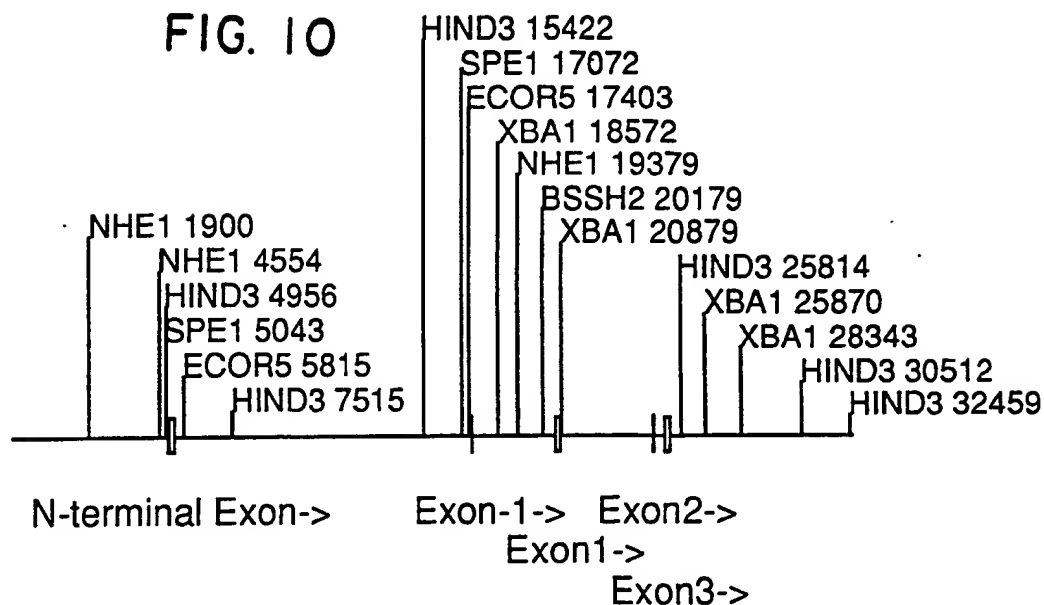
FIG. 12

AMINO TERMINAL EXON OF preproOIF

```
1   A AGC TTT AAA TAT TGC TTC GAT GGT CTG AAT TTT TAT
    T TCG AAA TTT ATA ACG AAG CTA CCA GAC TTA AAA ATA


      TTC CAG GGA AAA AGA GAG TTT TGT CCC ACA GTC AGC
      AAG GTC CCT TTT TCT CTC AAA ACA GGG TGT CAG TCG


      AGG CCA CTA GTT TAT TAA CTT CCA GTC ACC TTG ATT
      TCC GGT GAT CAA ATA ATT GAA GGT CAG TGG AAC TAA
              ▲
            SPE1


                          Signal Sequence
                        *** *** *** *** *** *** *** *** ***
                        Met Lys Thr Leu Gln Ser Thr Leu Leu
      TTT GCT AAA ATG AAG ACT CTG CAG TCT ACA CTT CTC
      AAA CGA TTT TAG TTC TGA GAC GTC AGA TGT GAA GAG
                              ▲
                            PST1


      *** *** *** *** *** *** *** *** *** ***
      Leu Leu Leu Leu Val Pro Leu Ile Lys Pro Ala Pro
      CTG TTA CTG CTT GTG CCT CTG ATA AAG CCA GCA CCA
      GAC AAT GAC GAA CAC GGA GAC TAT TTC GGT CGT GGT


      Pro Thr Gln Gln Asp Ser Arg Ile Ile Tyr Asp Tyr
      CCA ACC CAG CAG GAC TCA CG7 ATT ATC TAT GAT TAT
      GGT TGG GTC GTC CTG AGT GC0 TAA TAG ATA CTA ATA


      Gly Thr Asp Asn Phe Glu Glu Ser Ile Phe Ser Gln
      GGA ACA GAT AAT TTT GAA GAA TCC ATA TTT AGC CAA
      CCT TGT CTA TTA AAA CTT CTT AGG TAT AAA TCG GTT


      Asp Tyr Glu Asp Lys Tyr Leu Asp Gly Lys Ile Leu
      GAT TAT GAG GAT AAA TAC CTG GAT GGA AAA ATA TTA
      CTA ATA CTC CTA TTT ATG GAC CTA CCT TTT TAT AAT


      Arg Tyr Phe Ile Phe Tyr Ser Lys Phe Ser Phe Leu
      AGG TAC TTT ATT T8C TAT TCT AAA TTT AGC TT8 CTA
      TCC ATG AAA TAA A9G ATA AGA TTT AAA TCG AA9 GAT


      Asn Thr Ala
      AAT ACT GCC C
      TTA TGA CGG G
```

FIG.11-1

PART A

EXON -1

Human Genomic
TTAGTAATAGCATTCATAATAGCAATAATTTTACTACTAGAGGAGAAATATTTCAAACGATTGTTCATATT
********** ******* * ********* *** **** ***
TTAGTAATAACATTCATGATAGCAAGAATTTTAATACTAGAGGTCAGATATTTTAAATTATTG   ATT
Bovine Genomic

                                                              -90
YYYYYYYYYNYAG   ↓MetProThrCysLeuLeuCysValCysLeuSerGlySerValTyr
                                                              -100
TTTATGATTTTA TTCCCAATTTAGAAATGCCCACGTGTCTGCTGTGTTGTTAAGTGGCTCTGTATAC
********** ** ******** ****************************
TTTATGATTTTGTTTCCCAATTCAGAAATGCCCACATGCCTGCTATGTGTTGTTTAAGTGGCTCTGTATAC
MetIleLeuPheProAsnSerGluMetProThrCysLeuLeuCysValCysLeuSerGlySerValTyr
                                                              -90
                                                              -100

                           -80
CysGluGluValAspIleAspAlaValProProLysGluSerAlaTyrLeuTyrAlaArgPheAsn
TGTGAAGAAGTTGACATTGATGCTGTACCACCCTTACCAAAGGAATCAGCCTATCTTTACGCACGATTCAAC
********** ** ******************************* **
TGTGAAGAAGTTGACATTGATGCTGTACCACCTTTGCCAAAGGAATCAGCCTATCTTTATGCACGATTCAAC
CysGluGluValAspIleAspAlaValProProLysGluSerAlaTyrLeuTyrAlaArgPheAsn
                           -70
                           -80

FIG.11-2
PART B

```
                                                  JAGGTRAGT
                      -60                                  |
LysIleLysLysLeuThrAlaAlaLysAspPheAlaAspIleP|
AAAATTAAAAAGCTGACTGCCAAAGATTTTGCAGACATACGTAAGTTAAGCTGAAATATGATATTAGTATTA
*********** *** *************** *  ***** *  ***
AAAATTAAAAAGCTGACCGCCAAAGATTTTGCAGACATACGTAAGCTAATTTAAATAAAATATTAGTGTTA
LysIleLysLysLeuThrAlaAlaLysAspPheAlaAspIleP
                      -60
```

FIG.11-3
PART B

EXON 1

```
                                                    YYYYYYYYYYNYAG   -50
HUMAN                                                          ↓roAsnLeuArg
TCAGGTCCGATACAATAAGAATTTGCTTGGTAGTGGCTTTTATATTTTTCCCTCTAGCTAAGA
      *     * ** *    *    * **** ***  ****** ** *******
AATTTAACTTCATGTATTAAAAAATTTTTTTCTGTATTGTTGTATATTTTTCCTTCAAGCTAACTTAAGA
BOVINE                                                          roAsnLeuArg


                                    -30
ArgLeuAspPheThrGlyAsnLeuIleGluAspIleGluAspGlyThrPheSerLysLeuLeuLeuGlu
AGACTCGATTTTACAGGAAATTTGATAGAAGATATAGAAGATGGTACTTTTCAAAACTTTCTCTGTTAGAA
**** ******** ***** ****** ** ************ ****************
CGACTTGATTTTACGGGAAATTGATTGAAGACATAGAAGACGTACTTTTCAAAACTTTCTCTGTTAGAA
ArgLeuAspPheThrGlyAsnLeuIleGluAspIleGluAspGlyThrPheSerLysLeuLeuLeuGlu
                                    -40


                                    -10                                  1
GluLeuSerLeuAlaGluAsnGlnLeuLeuLysLeuProValLeuProProLysLeuThrLeuPheAsnAla
GAACTTCACTTGCTGAAAATCAACTACTAAAACTTCCAGTTCTTCCTCCCCAAGCTCACTTTATTTAATGCA
******* ***** ************* ** ************ **************
GAACTTACACTAGCTGAAAATCAAGTACTGAAGCTTCCCTCCCCAAGCTTACTTTATTTAATGCA
GluLeuThrLeuAlaGluAsnGlnLeuLeuLysLeuProValLeuProProLysLeuThrLeuPheAsnAla
                                    -20
```

FIG.11-4
PART C

JAGGTRAGT

```
                 10
LysTyrAsnLysIleLysSerArgGlyIleLysIleLysAlaAsnAlaPheLys
AAATACAACAAAATCAAAATCAAGAGTAGGGGAATCAAGCAAATGCATTCAAAGTAATTTTATAGTAAAACTAA
***** ***************** ***** **********************  *** *
AAATATAACAAAATCAAAGAGTAGAGGGAATCAAAGCAAATACATTCAAAGTAAGTATTTTATAGTATGACTGA
LysTyrAsnLysIleLysSerArgGlyIleLysIleLysAlaAsnThrPheLys


GCATAGTATGAGTGATGATGATAGATAAAGAGAAACTTTGGTCTAGA
*** ** **        *   ** *
GCACAATATATAAATATCATTGTTATACAGCCAATAATAACATGATAAAAG
```

FIG.11-5
PART C

EXON 2

```
                                                  CHO
                                                   |
                                LysLeuAsnAsnLeuThrPhe
HUMAN
TGAAAGAAATGTACAGAACCTCAATAACACTCACATGAAATTTCCTTACAGAAACTGAATAACCTCACCTTC
                      ***     *  *  *      ***** ***** *****
ATGCGCCTGATCTCAGGTATCAACTCAAGCAAAANCACCTTTTTCTTTACAGAAACTGCATAACCTCTCCTTC
BOVINE                          LysLeuHisAsnLeuSerPhe
                                                   |
                                                  CHO
YYYYYYYYYNYAG
```

```
LeuTyrLeuAspHisAsnAlaLeuGluSerValProLeuAsnLeuProGluSerLeuArgValIleHisLeu
CTCTACTTGGACCATAATGCCCTGGAATCCGTGCCCTCTTAATTTACCAGAAAGTCTACGGTGTAATTCATCTT
**********   **  *****      ********* *****************
CTCTACTTGGATCACAATGCTTTGGAATCTGTGCCTCTTAATTTACCAGAAAGTCTGCGGTGTAATTCATCTT
LeuTyrLeuAspHisAsnAlaLeuGluSerValProLeuAsnLeuProGluSerLeuArgValIleHisLeu
```

YYYYYYYYYNYAG

FIG.11-6
PART D

```
JAGGTRAGT
Gin
CAGGTATGATAGTTCCTT9CACAATATCTGACTCTAATTACCTATGAGTCAAAATCTTGTGA
********** * ****** * **** * ***** * *****      *    * *
CAGGTATGATTGCTCCTTCACAGTATCAGACTTTAATTCAAATTACTATGGGTCAATCTTA
Gin
JAGGTRAGT
```

FIG.11-7
PART D

EXON 3

```
                                                  50                                    60
                        YYYYYYYYYYNYAG            PheAsnAsnIleAlaSerIleThrAspAspThrPheCys
HUMAN
TTTCTGAACAATGTTTTCCCTCGTTTTTTTCAGTTCAACAACATAGCTTCAATTACAGATGACACATTCTGC
  *  * *  * **  * *** *** *  *** ** **********  **********************  *** **
GGAGTTTATAAAGTATTTTCCCTTGTTTCCAGTTTAACAACATAACTTCAATTACAGATGATACATTCTGC
                        YYYYYYYYYYNYAG            PheAsnAsnIleThrSerIleThrAspAspThrPheCys
BOVINE
                                            |
                                            CHO
```

```
                                                  70                                    80
         |        LysAlaAsnAspThrSerTyrIleArgAspArgIleGluGluIleArgLeuGluGlyAsnProIleValLeu
         CHO
AAGGCTAATGACAACCAGTTACATCCGGGACCGCATTGAAGAGATACGGCCTGGAGGCAATCCAATCGTCCTG
*****************  ** *****************  * ***********  **  ** *** ****
AAGGCTAATGACACACCAGTTATATTCGGGACCGCATTATTGAAGAGATACGCTTGGAGGGCAACCCCGTCATCCTG
         |        LysAlaAsnAspThrSerTyrIleArgAspArgIleGluGluIleArgLeuGluGlyAsnProValIleLeu
         CHO
```

FIG.11-8
PART E

```
                                              100
GlyLysHisProAsnSerPheIleCysLeuLysArgLeuProIleGlySerTyrPhe
GGAAAGCATCCAAACAGTTTATTTGCTTAAAAAGATTACCGATACGGTCATACTTTAACCTCTATTGGTA
** ** ****** *********** ** *** *** ******** * **  *
GGGAAACATCCCAACAGTTTATCTGCTTAAAAAGATTGCCTATAGGGTCATACATTTAACCACNATCAATG
GlyLysHisProAsnSerPheIleCysLeuLysArgLeuProIleGlySerTyrIle


CAACATATAAATGAAAGTACACCTAACACTAATAGTCTGTCTCAACAATGTGTAAAGGAACTTAAGTATTGGT
**  ***        *  ***  *** ********** ****  * *  * ****
CAG ATA    GCTAAAGTACACACATACTTATAATCTGTCTCAACAATGTCTCAAACCAGCATAAATATT

      Poly A
      Addition
       Site
TTAATATTAACCTTGTATCTCATT
*********  *** **** * *
TAATATTAATTTTGCATCTCTGACT
     Poly A
     Addition
      Site
```
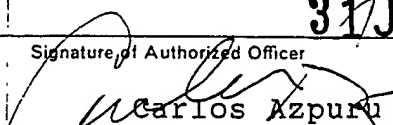
FIG.11-9
PART E

INTERNATIONAL SEARCH REPORT

International Application No. **PCT/US90/04745**

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) 6

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC (5): C07K 15/06,3/02,3/20,13/00; A61K 35/12,37/12;
A01N 63/02; A23J 1/10 U.S. CL. 424/95

II. FIELDS SEARCHED

	Minimum Documentation Searched 7	
Classification System	Classification Symbols
U.S.	424/95; 514/2, 12, 21; 530/350, 353, 355, 356 530/ 414, 416, 417, 540

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched 8

III. DOCUMENTS CONSIDERED TO BE RELEVANT 9

Category *	Citation of Document, 11 with indication, where appropriate, of the relevant passages 12	Relevant to Claim No. 13
A	US, A, 4,294,753 (URIST) 13 OCTOBER 1981 See entire document.	1-7
A	US, A, 4,434,094 (SEYEDIN ET AL) 28 FEBRUARY 1984; See entire document.	1-7
A	US, A, 4,455,256 (URIST) 19 JUNE 1984 See entire document.	1-7
A	US, A, 4,608,199 (CAPLAN ET AL) 26 AUGUST 1986; See entire document.	1-7
A	US, A, 4,627,982 (SEYEDIN ET AL) 09 DECEMBER 1986; See entire document.	1-7
A	US, A, 4,774,322 (SEYEDIN ET AL) 27 SEPTEMBER 1988; See entire document.	1-7
A	US, A, 4,789,732 (URIST) 06 DECEMBER 1988 See entire document.	1-7
A	US, A, 4,795,804 (URIST) 03 JANUARY 1989 See entire document.	1-7

* Special categories of cited documents: 10

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search	Date of Mailing of this International Search Report
19 DECEMBER 1990	3 17 JAN 1991
International Searching Authority	Signature of Authorized Officer
ISA/US	Carlos Azpuru

Form PCT/ISA/210 (second sheet) (Rev.11-87)

III. DOCUMENTS CONSIDERED TO BE RELEVANT   (CONTINUED FROM THE SECOND SHEET)

Category *	Citation of Document, with indication, where appropriate, of the relevant passages	Relevant to Claim No
A	US, A, 4,810,691 (SEYEDIN ET AL) 07 MARCH 1989; See entire document.	1-7
A	US, A, 4,816,442 (McPHERSON ET AL) 28 MARCH 1989; See entire document.	1-7
A	WO, A1, WO88/00205 (WANG ET AL) 14 JANUARY 1988; See entire document.	1-7
A	EP, A2, 0212474 (URIST ET AL) 04 MARCH 1987 See entire document.	1-7
A	Reddi et al., Proc. Nat. Acad. Sci. USA (1972) Vol. 69: 1601-1605. See entire document.	1-7
A	Urist et al., Scienu (1965) Vol. 150: 893-899. See entire document.	1-7
A	Urist et al., Clin. Orthop. (1968) Vol. 56: 37-50. See entire document.	1-7
A	Urist et al., Clin. Orthop. Rel. Res. (1982) Vol. 162: 219-232. See entire document.	1-7
A	Urist et al., Science (1983) Vol. 220: 680-685. See entire document.	1-7
A	Urist et al., Proc. Nat. Acad. Sci. USA (1984) Vol. 81: 371-375. See entire document.	1-7
A	EP, A3, 0128041 (BAYLINK) 12 DECEMBER 1984 See entire document.	1-7
A	Jennings et al., Meth. Enzymol. (1987) Vol. 146: 281-283. See entire document.	1-7

Form PCT/ISA/210 (extra sheet) (Rev.11-87)

FURTHER INFORMATION CONTINUED FROM THE SECOND SHEET

V.☐ OBSERVATIONS WHERE CERTAIN CLAIMS WERE FOUND UNSEARCHABLE [1]

This international search report has not been established in respect of certain claims under Article 17(2) (a) for the following reasons:

1.☐ Claim numbers , because they relate to subject matter [12] not required to be searched by this Authority, namely:

2.☐ Claim numbers , because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out [13], specifically:

3.☐ Claim numbers_____, because they are dependent claims not drafted in accordance with the second and third sentences of PCT Rule 6.4(a).

VI.☒ OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING [2]

This International Searching Authority found multiple inventions in this international application as follows:

```
Group I: Claims 1-7, drawn to a polypeptide.
Group II: Claim 8, drawn to an antibody.
Group III: Claims 9 & 10, drawn to a recombinant polynucleotide.
Group IV: Claims 11-14, drawn to a recombinant vector and host
          cell.
```

1.☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims of the international application.

2.☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims of the international application for which fees were paid, specifically claims:

3.☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claim numbers:

   1-7

4.☐ As all searchable claims could be searched without effort justifying an additional fee, the International Searching Authority did not invite payment of any additional fee.

Remark on Protest

   ☐ The additional search fees were accompanied by applicant's protest.

   ☐ No protest accompanied the payment of additional search fees.

Form PCT/ISA/210 (supplemental sheet (2) (Rev. 11-87)